

[illegible]

REFERENCE	TITLE	JOURNAL	COMMENT
1 (bases 1 to 529)	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .		
	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
	Tumor Gene Index		
	Unpublished (1997)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgaps-internal.nih.gov		
	CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo		
	, Ph.D.		
	CDNA Library Arrayed by: Greg Lennon, Ph.D.		
	DNA Sequencing by: Washington University Genome Sequencing Center		
	Clone distribution: NCI-CGAP clone distribution information can be		
	found through the I.M.A.G.E. Consortium/BLM at:		
	<a href="http://www.bio.llnl.gov/bdrip/image/image.html">www.bio.llnl.gov/bdrip/image/image.html</a>		
	Insert Length: 1076 Std Error: 0.00		
	Seq primer: -40ml3 fwd. RT from Amersham		
	High quality sequence stop: 411.		
FEATURES			
SOURCE			
1..529	Location/Qualifiers		
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone="IMAGE:1641675"			
/clone_lib="Soares_testis_NHT"			
/sex="male"			
/lab_host="DH10B"			
/note="Vector: pT773D-Pac (Pharmacia) with a modified			
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA			
was prepared from mRNA obtained from Clontech Laboratories			
, Inc., and primed with a Not I - oligo(dT) primer [5'			
TGTTACCAATCGTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].			
Double-stranded cDNA was ligated to Eco RI adaptors			
(Pharmacia), digested with Not I and cloned into the Not I			
and Eco RI sites of the modified pT773 vector. Library			
went through one round of normalization to Cot1, and was			
constructed by Bento Soares and M. Fatima Bonaldo."			
BASE COUNT	80 a 143 c 134 g 172 t		
ORIGIN			
Query Match	27.9%; Score 529; DB 10; Length 529;		
Best Local Similarity	100.0%; Pred. No. 2.9e-104;		
Matches 529; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
454 aagagagatgaagcttcacgtcagtcgaagctcaccacacagatgacgtcccatccacc	513		
529 aaggaatgaaaccttcacgtcgaagctcaccacacagatgacgtcccatccacc	470		
514 cacttcacagctgacagaaagcagacgttcacgtccgtgctgagaggtcagcaaac	573		
469 cacttcacaggtgacagaaagcagacgttcacgtccgtgctgagaggtcagcaaac	410		
574 gtggaagagctcctaaacctcctctgtctccctgaggagcagagagacgccaagac	633		
409 gtggaagagctcctaaacctcctctgtctccctgaggagcagagagacgccaagac	350		
634 aagcagagcaagagtgtagcagcagcagagagcagcagcaagaagaagaaga	693		
349 aagcagagcaagagtgtagcagcagcagagagcagcagcaagaagaagaaga	290		
694 gggcagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga	753		
289 gggcagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga	230		
754 cagggagacgaagaagaagaagaagaagaagaagaagaagaagaagaaga	813		
229 cagggagacgaagaagaagaagaagaagaagaagaagaagaagaagaaga	170		
814 ttccactctgaactctatctctttaaaccctctctctttgtctcccggttacagaagta	873		
169 ttccactctgaactctatctctttaaaccctctctctttgtctcccggttacagaagta	110		
874 gagtctactcctatgataatgagagaacatccagagagctatttgatccagcccaagaata	933		

DB 109 GAGCTACTCTCATGTAATNGAGAAATCCAGAGCTCATTCATGACCCAGAAATA 50  
 QY 934 gatgaatgaatgaatataatgatgaactcctactctgagaaccaa 982  
 DB 49 GATGAATGAATGAATATATATATGATGAACTCTACTGAGAAACCAA 1

RESULT 7  
 AA424694 496 bp mRNA EST 16-OCT-1997  
 LOCUS zulu10.s1 Soares\_testis.NHT Homo sapiens cDNA clone IMAGE:731779  
 DEFINITION 3' similar to TR:G475021 G475021 SP32 PRECURSOR ; contains element  
 TARI repetitive element ; mRNA sequence.

ACCESSION AA424694  
 VERSION AA424694  
 KEYWORDS GI:2102744  
 SOURCE EST.  
 ORGANISM human.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 496)  
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Giesel, G., Jost, S.,  
 Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,  
 Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie,  
 T., Waterston, R. and Wilson, R.  
 'T', Waterston, R. and Wilson, R.  
 Mashu-Merck EST Project 1997  
 Unpublished (1997)  
 CONTACT: Wilson R.  
 TITLE Washington University School of Medicine  
 JOURNAL 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 COMMENT Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Possible reversed clone: similarity on wrong strand  
 Seq primer: -41m13 fwd. ET from Amersham  
 High quality sequence stop: 313.

## FEATURES

Location/Qualifiers  
 1..496  
 /organism="Homo sapiens"  
 /db\_xref="GDB:5928879"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:731779"  
 /clone\_1lb="Soares\_testis\_NHT"  
 /sex="male"  
 /lab\_host="DH10B"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
 was prepared from mRNA obtained from Clontech Laboratories  
 , Inc., and primed with a Not I - oligo(dT) primer [5'  
 TGTACCAATCGAAGTGGAGCGGCCCAATTTTCTTTT 3']  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization to Cot5, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."  
 BASE COUNT 75 a 134 c 127 g 160 t  
 ORIGIN

Query Match 26.0%; Score 492.8; DB 10; Length 496;  
 Best Local Similarity 99.6%; Pred. No. 1.9e-96;  
 Matches 494; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 487 accaagatgaactcccatcaccatcaccatcagaagacagcgaacttcag 546  
 DB 496 ACCAGATGACCTCCCTCCATCTCACCCACCTTCACAGAGAGAGCCACCTTCAG 437  
 547 ccctggccctgagaggtcagcacaacgtggaagagctcctacacactcctctg 606  
 36 CCCTGGCCTGAGAGGCTCAGCAACAGTGGAGAGCTCTTACATCTCTCTTGTCCCTG 377

QY 607 ggaagccaagagcaagcgccagagcacaagcaggaagagtgagcagcaggaag 666  
 DB 376 GGAAGCCAGGAGCAAGGCCAGACAGACAGAGAGAGAGAGGACATGACAGAG 317  
 QY 667 ccgacacaagacacaagcaggaagagggcagaaacaggaagagcagaaggaacag 726  
 DB 316 CCGACACAGACACACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 257  
 QY 727 gaagaggaaggaagcaggaaggaaggaaggaaggaaggaaggaaggaaggaag 786  
 DB 256 GAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 197  
 QY 787 cagctgcaggaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 846  
 DB 196 CAGCTGCAGACAGACTCAGAGCCCAAGTTCTACTGAAATCTATCTTCTTACCTTCC 137  
 QY 847 tctttgtcccggtacgagagtagtagtctactcctatgataatgagaacatcag 906  
 DB 136 TCTTTTGTCTCCCGGTRACGAGAGTAGTACTCTCTATGATGAGAGACATCCAG 77  
 QY 907 gagctcattcagatcagccaggaatagatgaatgaatgaatgaatgaatgaatc 966  
 DB 76 GAGCTCATTCATCAGCCAGAGAAATGATGAATGAATGATGATGATGATGATGATC 17  
 QY 967 tactgagaagaccaa 982  
 DB 16 TACTGAGAGAAACCAAA 1

RESULT 8  
 AA824340 496 bp mRNA EST 31-DEC-1998  
 LOCUS a30e09.s1 Soares\_testis\_NHT Homo sapiens cDNA clone 1391848 3'  
 DEFINITION similar to FR:060485 060485 SP32 PRECURSOR ; mRNA sequence.  
 ACCESSION AA824340  
 VERSION AA824340  
 KEYWORDS GI:2896410  
 SOURCE EST.  
 ORGANISM human.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 496)  
 NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Contact: cga@bs-riemail.nih.gov  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo  
 , Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CCAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 www.bio.llnl.gov/bbrp/image/image.html  
 Insert Length: 609 Std Error: 0.00  
 Seq primer: -40m13 fwd. ET from Amersham.  
 FEATURES  
 Location/Qualifiers  
 1..496  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="1391848"  
 /clone\_1lb="Soares\_testis\_NHT"  
 /sex="male"  
 /lab\_host="DH10B"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
 was prepared from mRNA obtained from Clontech Laboratories  
 , Inc., and primed with a Not I - oligo(dT) primer [5'  
 TGTACCAATCGAAGTGGAGCGGCCCAATTTTCTTTT 3']  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT73 vector. Library

to identify individuals or in diagnostic procedures to identify individuals having genetic diseases resulting from abnormal expression of the genes corresponding to the extended cDNAs. They are also useful for constructing a high resolution map of the human chromosomes. They can also be used for gene therapy to control or treat genetic diseases.

Sequence 964 BP; 258 A; 294 C; 227 G; 180 T; 5 other:

Query Match 49.4%; Score 937; DB 20; Length 964;  
Best Local Similarity 99.3%; Pred. No. 6,3e-235;  
Matches 948; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

```

30 gggcgagatcttcggcgatgagaaagcagcgtgtcttcctcctcactcgtgaa 89
   |||
2 gggcgagatcttcggcgatgagaaagcagcgtgtcttcctcctcactcgtgaa 61
   |||
90 ggtgctctctgctctctgacacgtgcgacgacgagattcagacgacccactcc 149
   |||
62 ggtgctctctgctctctgacacgtgcgacgacgagattcagacgacccactcc 121
   |||
150 aggcagcctctctcctcagcagcagcagcagcagcagcagcagcagcagcag 209
   |||
122 aggcagcctctctcctcagcagcagcagcagcagcagcagcagcagcagcag 181
   |||
210 gaagcagagactactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 269
   |||
182 gaagcagagactactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 241
   |||
270 ccagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 329
   |||
242 ccagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 301
   |||
330 ccctatgctcctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 389
   |||
302 ccctatgctcctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 361
   |||
390 cgtctactatgacagagagctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 449
   |||
362 cgtctactatgacagagagctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 421
   |||
450 tctcaagagatagaagctcagctcagctcagctcagctcagctcagctcagctc 508
   |||
422 tctcaagagatagaagctcagctcagctcagctcagctcagctcagctcagctc 481
   |||
509 caccacactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 568
   |||
482 caccacactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 541
   |||
69 acaacgtgagagagctcctcctcctcctcctcctcctcctcctcctcctcctc 628
   |||
542 acaacgtgagagagctcctcctcctcctcctcctcctcctcctcctcctcctc 601
   |||
659 agcacaagcagagcagagcagagcagagcagagcagagcagagcagagcagagc 688
   |||
602 agcacaagcagagcagagcagagcagagcagagcagagcagagcagagcagagc 661
   |||
669 aagagggcagagcagagcagagcagagcagagcagagcagagcagagcagagc 748
   |||
662 aagagggcagagcagagcagagcagagcagagcagagcagagcagagcagagc 721
   |||
749 aagagcagagcagagcagagcagagcagagcagagcagagcagagcagagcagag 808
   |||
722 aagagcagagcagagcagagcagagcagagcagagcagagcagagcagagcagag 781
   |||
809 ccaagttcactctgaatctatcttcttaacccctcctcttctgctcccggtacag 868
   |||
782 ccaagttcactctgaatctatcttcttaacccctcctcttctgctcccggtacag 841
   |||
869 aagtagagctactcctctatgataagagagacatccagagcattcagatccagcag 928
   |||
842 aagtagagctactcctctatgataagagagacatccagagcattcagatccagcag 901

```

QY 929 aaatagatgaatgaatgaatataatagatgagaactcctactgagagacccaaa 983  
DB 902 aaatagatgaatgaatgaatataatagatgagaactcctactgagagacccaaa 956

RESULT 7  
AAV88163  
AAV88163 standard; cDNA; 308 BP.

AC AAV88163;  
XX  
DT 12-FEB-1999 (first entry)  
XX  
XX

DE EST clone GA63.

KW Expressed sequence tag; secreted protein; haematopoiesis regulator;  
KW tissue growth; actinin; inhibin; tumour invasion suppressor; EST; human;  
KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;  
KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.  
XX  
XX

OS Homo sapiens.

PN WO9845437-A2.

XX 15-OCT-1998.

PF 10-APR-1998; 98MO-US06956.

XX 10-APR-1997; 97US-0837312.

PR (GENE) GENETICS INST INC.

XX Agostino M, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

PI Racine LA, Spaulding V, Treacy M;

XX MPI; 1999-070078/06.

PT New polynucleotides encoding human secreted proteins - derived from  
PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,  
PT ovary, pituitary, retina and colon cDNA libraries

XX Claim 1; Page 304; 641pp; English.

XX The present sequence represents an expressed sequence tag (EST), and is  
XX a polynucleotide of the invention. The polynucleotides of the invention  
XX are all secreted EST sequences and proteins isolated from a variety of human tissue  
XX sources. The EST sequences and proteins encoded by them are predicted to  
XX have useful biological activities which would make them suitable for  
XX treating, preventing or ameliorating medical conditions in humans and  
XX animals, although no supporting data is given. Suggested activities  
XX include nutritional activity, immune stimulating or suppressing activity,  
XX haematopoiesis regulating activity, tissue growth activity,  
XX activin/inhibin activity, chemotactic/chemokine activity, haemostatic  
XX and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
XX activity, cadherin/tumour invasion suppressor activity, tumour inhibition  
XX activity. The EST sequences are also stated to be useful for gene  
XX therapy.

XX Sequence 308 BP; 85 A; 100 C; 75 G; 48 T; 0 other;

Query Match 14.9%; Score 282.4; DB 20; Length 308;  
Best Local Similarity 96.3%; Pred. No. 3.8e-64;  
Matches 289; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

QY 395 actatgcaagaagagctctgtgttcacgacgagctctatctcctaactaactca 454
   |||
DB 3 atcggcttcataagcctagagctccagcagcagctctatctcctaactaactca 62
   |||
QY 455 aggaatagaagctcagctgaagctcaccacacagatgacctcccatcaccoccc 514
   |||
DB 63 aggaatagaagctcagctgaagctcaccacacagatgacctcccatcaccoccc 122

```

